

УДК 53.05: 617.753

Павло Тимків, Любомир Демчук

Тернопільський національний технічний університет імені Івана Пулюя, Україна

**ROC-АНАЛІЗ МЕТОДІВ ПАРАМЕТРИЧНОЇ ІДЕНТИФІКАЦІЇ
ОБЧИСЛЮВАЛЬНОЇ МОДЕЛІ НИЗЬКОІНТЕНСИВНОГО
ЕЛЕКТРОРЕТИНОСИГНАЛУ**

Відомі роботи в яких запропоноване застосування фільтру Калмана та методи параметричної ідентифікації обчислювальної моделі низькоінтенсивного електроретиносигналу. Проте використаний метод прямого направленого перебору (пошуку) коефіцієнтів, має значну часову складність, що унеможливорює автоматизоване застосування і переналаштування фільтру. Тому використовують удосконалений метод визначення коефіцієнтів математичної моделі низькоінтенсивного електроретиносигналу шляхом перебору у декілька ітерацій зі зміною кроку перебору. Виникає необхідність оцінювання удосконаленого методу параметричної ідентифікації на основі ROC-аналізу.

Ключові слова: електроретиносигнал, низька інтенсивність, фільтр Калмана, часова складність, ROC-аналіз, AUC, GINI.

Pavlo Tymkiv, Liubomyr Demchuk

**ROC-ANALYSIS OF METHODS PARAMETRIC IDENTIFICATION OF THE
COMPUTATIONAL MODEL LOW-INTENSIVE ELECTRORETINOSIGNAL**

The ROC-analysis of the advanced method for determining the coefficients of the mathematical model of low-intensity electrodynamics was conducted, and the AUC and GINI index were determined.

Keywords: electroretinosignal low intensity, Kalman filter, time complexity, ROC-analysis, AUC, GINI

Відомі роботи, в яких на евристичних засадах обґрунтовано застосування фільтру Калмана та використано метод параметричної ідентифікації обчислювальної моделі відгуку ретини (електроретиносигналу (ЕРС)) на світлове подразнення (метод-прототип).

При цьому використовується метод послідовного підбору коефіцієнтів обчислювальної моделі, якому властива значна часова складність. Для зменшення часу підбору використовують метод параметричної ідентифікації у декілька ітерацій зі зміною кроку перебору (удосконалений метод). Тому виникає потреба у застосуванні ROC-аналізу і визначення показників AUC та індексу GINI для оцінювання удосконаленого методу.

Шляхом імітаційного моделювання отримано статистику низькоінтенсивних ЕРС $x(t)$ з нормальним розподілом ймовірності (рис.1).

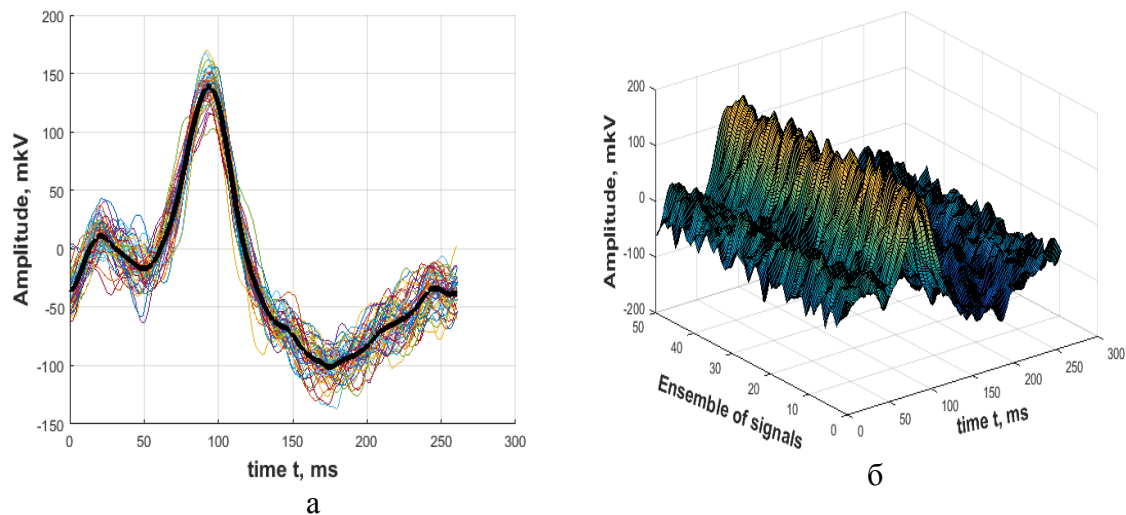


Рис. 1. Ансамбль імітаційно-модельованих низькоінтенсивних ЕРС:

- а) – ансамбль модельованих низькоінтенсивних ЕРС методом-прототипом;
- б) – ансамбль модельованих низькоінтенсивних ЕРС удосконаленим методом

Густини розподілу ймовірностей середньоквадратичної похибки (СКП) моделювання як міри оцінювання методу визначення коефіцієнтів математичної моделі низькоінтенсивного ЕРС, отриманих тестових статистик, наведено на рис.2.

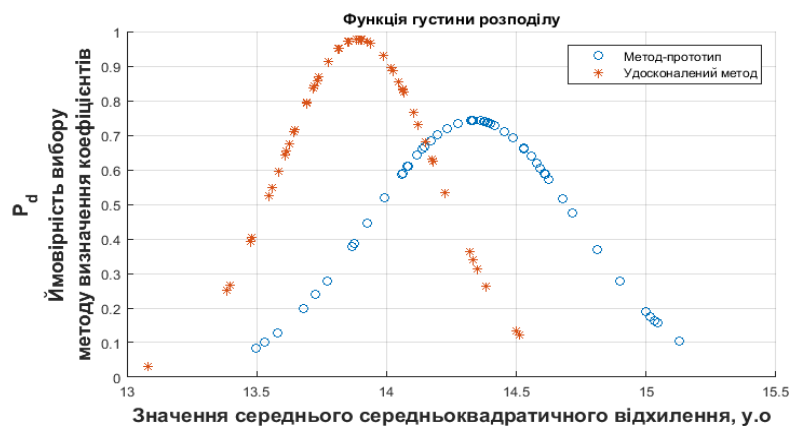


Рис. 2. Функція густини розподілу СКП методу-прототипу та удосконаленого методу

Задача вибору методу визначення коефіцієнтів математичної моделі низькоінтенсивного ЕРС при опрацюванні фільтром Калмана вирішується методами статистичної теорії вибору рішень.

При цьому умовна ймовірність помилкового визначення P_{fa} , трактується як ймовірність того, що значення тестової статистики $q(t)$ перевищить деяке граничне значення порогу γ , що призведе до помилкового рішення.

Відповідно до критерію Неймана-Пірсона ймовірність помилкового вибору методу P_{fa} повинна задаватися:

$$P_{fa} = \int_{\gamma}^{\infty} W_0(q) dq = 1 - \Phi\left(\frac{\gamma - m_0}{\sqrt{D_0}}\right) \quad (1)$$

де $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{u^2}{2}} du$ — інтеграл ймовірності.

Тоді ймовірність правильного вибору методу:

$$P_d = \int_{\gamma}^{\infty} W_1(q) dq = 1 - \Phi\left(\frac{\gamma - m_1}{\sqrt{D_1}}\right) \quad (2)$$

Характеристики достовірності вибору методу визначення коефіцієнтів є залежністю ймовірності правильного визначення P_d від коефіцієнту девіації, при фіксованій ймовірності помилкового вибору методу P_{fa} . ($P_{fa} = 0,1$; $P_{fa} = 0,01$; $P_{fa} = 0,001$; $P_{fa} = 0,0001$). Відповідні графіки наведено на рис.3

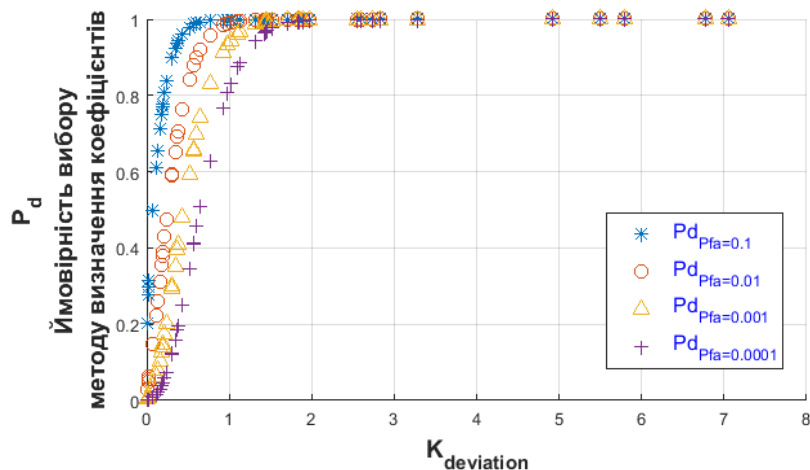


Рис. 3 Функціональна залежність ймовірності вибору методу від коефіцієнта девіації

Базуючись на результатах бінарної класифікації адаптованого критерію Неймана-Пірсона, та обчислених значеннях ймовірності правильного вибору методу P_d при заданні ймовірності хибного вибору P_{fa} , в подальшому дослідженні, необхідно визначити залежність кількості вірно класифікованих позитивних результатів від кількості невірно класифікованих помилкових результатів за допомогою ROC-аналізу.

ROC-характеристика методів визначення коефіцієнтів математичної моделі низькоінтенсивного ЕРС при опрацюванні фільтром Калмана для методу-прототипу та удосконаленого методу набуває вигляду (рис. 4).

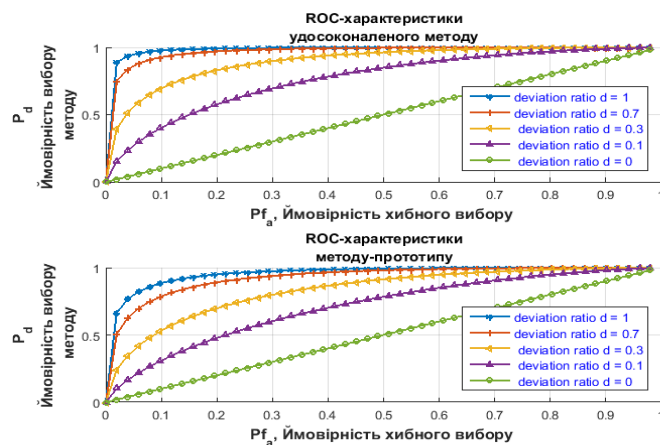


Рис. 4. ROC-криві методу-прототипу та удосконаленого методу

Чим ближче ROC-крива до верхнього лівого кута, тим метод є ефективнішим. І навпаки, чим ближче вона до діагоналі, тим менш ефективний метод. Проте візуальне порівняння ROC-кривих не завжди дозволяє виявити найбільш ефективний метод. Для кількісної інтерпретації ROC-аналізу визначимо показник AUC (Area Under ROC Curve) – площа, обмежена ROC-кривою і віссю частки помилкових позитивних класифікацій.

Метод-прототип		Удосконалений метод	
Deviation ratio (d)	Показник AUC	Deviation ratio (d)	Показник AUC
$d_1 = 1$	$AUC_1 = 0,9365$	$d_1 = 1$	$AUC_1 = 0,9628$
$d_2 = 0,7$	$AUC_2 = 0,9065$	$d_2 = 0,7$	$AUC_2 = 0,9486$
$d_3 = 0,3$	$AUC_3 = 0,8111$	$d_3 = 0,3$	$AUC_3 = 0,8757$
$d_4 = 0,1$	$AUC_4 = 0,6904$	$d_4 = 0,1$	$AUC_4 = 0,7474$

Після визначення значення AUC обчислимо індекс GINI (G) за формулою:

$$G_{(GINI)} = 2 \cdot (AUC - 0,5) \quad (3)$$

де $G_{(GINI)}$ – якість ROC-кривої.

$$G_{1(GINI)} = 0,9256, G_{2(GINI)} = 0,8972, G_{3(GINI)} = 0,7514, G_{4(GINI)} = 0,4948.$$

Максимальне значення індекса $G_{1(GINI)}$ дорівнює 0,9256.

Висновки. Проведений ROC-аналіз показав вищу достовірність запропонованого удосконаленого методу визначення коефіцієнтів, а визначення кількісного показника AUC та індексу $G_{(GINI)}$ ($AUC_1 = 0,9628$, $AUC_2 = 0,9486$, $AUC_3 = 0,8757$, $AUC_4 = 0,7474$ та $G_{1(GINI)} = 0,9256$, $G_{2(GINI)} = 0,8972$, $G_{3(GINI)} = 0,7514$, $G_{4(GINI)} = 0,4948$ відповідно) для двох методів – більшу повноту та чутливість удосконаленого методу в порівнянні з методом прототипом.